(51) International Patent Classification 6 :		(11) International Publication Number:	**WO 98/05066**
H01L 21/66	A2	(43) International Publication Date:	5 February 1998 (05.02.98)

(54) Title: METHODS AND APPARATUS FOR THE IN-PROCESS DETECTION AND MEASUREMENT OF THIN FILM LAYERS

(57) Abstract

The present invention provides methods and apparatus which permits in-process, in-situ, substantially real time measurement of the actual thickness of a surface layer of a work piece, e.g., a semiconductor wafer. A probe (500) disposed proximate to the outer perimeter of a polishing pad (126) on a CMP table, such that the probe establishes optical contact with the wafer surface (304) as a portion of the wafer extends beyond the outer perimeter (330) of the polishing pad (126). A nozzle (312) may be provided to apply a stream of compressed air at the disk surface under inspection to thereby remove excess slurry from the local region of the workpiece being inspected. A broad band light source (322) is employed in conjunction with a fiber optic cable (318) to direct light at the wafer surface. A bifurcated probe is employed such that the light applied to the workpiece surface is reflected back to and captured by a corresponding optical sensor connected to a fiber optic cable (320). The captured reflected light received by the receptor sensor and fiber optic cable assembly is applied to a photospectrum meter (324) which analyzes the reflected light. An output signal from the photospectrum meter is transmitted to a processor (326) which includes a smart algorithm configured to calculate the thickness of the surface layer. Alternatively, the reflective characteristics of the semiconductor layers may affect the nature of the reflected signal; changes in the reflected signal can be detected to indicate when a metallic layer has been removed from an oxide layer.

METHODS AND APPARATUS FOR THE IN-PROCESS
DETECTION AND MEASUREMENT OF THIN FILM LAYERS

Technical Field

The present invention relates, generally, to methods and apparatus for measuring the

5   thickness of different material layers on a semiconductor workpiece during the planarization

process and, more particularly, to a technique for transmitting a broad band light source at the

surface of the workpiece to be measured, and analyzing the reflection of the light source to

obtain real time layer thickness information.

Background Art and Technical Problems

10   The production of integrated circuits begins with the creation of high-quality

semiconductor wafers. During the wafer fabrication process, the wafers may undergo multiple

masking, etching, and dielectric and conductor deposition processes. Because of the high-

precision required in the production of these integrated circuits, an extremely flat surface is

generally needed on at least one side of the semiconductor wafer to ensure proper accuracy and

15   performance of the microelectronic structures being created on the wafer surface. As the size

of the integrated circuits continues to decrease and the density of microstructures on an

integrated circuit increases, the need for precise wafer surfaces becomes more important.

Therefore, between each processing step, it is usually necessary to polish or planarize the

surface of the wafer to obtain the flattest surface possible.

20   For a discussion of chemical mechanical planarization (CMP) processes and

apparatus, see, for example, *Arai, et al.*, U.S. Patent No. 4,805,348, issued February, 1989;

Arai, et al., U.S. Patent No. 5,099,614, issued March, 1992; *Karlsrud et al.*, U.S. Patent No.

5,329,732, issued July, 1994; *Karlsrud*, U.S. Patent No. 5,498,196, issued March, 1996; and

Karlsrud et al., U.S. Patent No. 5,498,199, issued March, 1996.

25   Such polishing is well known in the art and generally includes attaching one side of

the wafer to a flat surface of a wafer carrier or chuck and pressing the other side of the wafer

against a flat polishing surface. In general, the polishing surface comprises a horizontal

polishing pad that has an exposed abrasive surface of, for example, cerium oxide, aluminum

oxide, fumed/precipitated silica or other particulate abrasives. Polishing pads can be formed of

30   various materials, as is known in the art, and which are available commercially. Typically, the

polishing pad may be a blown polyurethane, such as the IC and GS series of polishing pads

available from Rodel Products Corporation in Scottsdale, Arizona. The hardness and density

of the polishing pad depends on the material that is to be polished.

During the polishing or planarization process, the workpiece or wafer is typically pressed against the polishing pad surface while the pad rotates about its vertical axis. In addition, to improve the polishing effectiveness, the wafer may also be rotated about its vertical axis and oscillated back and forth over the surface of the polishing pad. It is well known that

5   polishing pads tend to wear unevenly during the polishing operation, causing surface
•   irregularities to develop on the pad. To ensure consistent and accurate planarization and polishing of all workpieces, these irregularities must be removed.

A well prepared polishing pad facilitates the uniform, high-precision planarization of workpieces. This is particularly important when polishing down the oxide and metalic layers

10   on a semiconductor wafer during the manufacture of integrated circuit chips.

Presently known methods for measuring the thickness of an oxide layer on a semiconductor wafer involve measuring the total thickness of an applied oxide layer, determining the desired thickness of the oxide layer after planarization, calculating the pressure to be applied during the polishing or planarization process, and further calculating the

15   approximate time required to remove a predetermined amount of oxide layer for a given pressure and slurry combination. Once the desired removal rate (often expressed in nanometers per minute) is ascertained, a statistical inference is employed to determine the approximate amount of time necessary to remove a desired amount of material. After the wafers have undergone planarization for an amount of time calculated to remove a desired thickness of the oxide layer,

20   the workpieces are removed from the machine and the actual thickness of the oxide layer is measured, for example, through the use of laser interferometric techniques. If it is determined that the oxide layer is still too thick after initial planarization, the workpieces must be reinstalled onto the CMP machine for further oxide layer removal. If, on the other hand, an excessive amount of oxide layer has been removed, it may be necessary to scrap the disks, resulting in

25   substantial unnecessary costs.

Further the methods of calculating oxide layer thicknesses currently known in the art are only useful for non-patterned wafers, and generally do not work on wafers having a substantially repeating surface pattern.

Other techniques may be employed to determine when the layer of tungsten, or other

30   metallic material, has been removed from the oxide layer. Prior art methodologies detect the change in polishing pad motor current, which typically changes in response to the exposure of different semiconductor layers. However, this technique is of limited utility where different

2

slurries and other consumable sets are used during the polishing process. In particular, false trigger may result from an increase or decrease in the surface friction caused by a change in the slurry characteristics rather than by a change in the exposed layer composition. False triggers may also be caused by noise inherent in the polishing system.

5   Accordingly, a technique is needed which accurately measures the oxide layer (and particularly the end point) thickness which overcomes the shortcoming of the prior art. In addition, a technique is needed which accurately detects the removal of a conductor or metalic layer from the oxide substrate layer.

Summary of the Invention

10   Accordingly, it is an advantage of the present invention that improved methods and apparatus for measuring the thickness of layers on workpiece surfaces are provided.

Another advantage of the present invention is that it facilitates the in-process, in-situ, substantially real time measurement of the actual thickness of a surface layer of a workpiece under inspection, for example, a semiconductor wafer (either patterned or non-patterned), or the
15   like.

A further advantage is that the present invention employs a smart algorithm configured to calculate the thickness of the oxide layer from information gathered from light signals reflected from the surface of the wafer.

Another advantage of the present invention is that the oxide layer thickness as a
20   function of time is displayed on a view screen for convenient observation by the operator of the machine. Additional functionality may be incorporated into the present invention to enable it to accurately predict the amount of time remaining and the planarization pressure needed to achieve an optimum end point oxide layer thickness.

A further advantage is that the present invention may be alternatively configured to
25   optically detect the endpoint when a metallic layer is removed from the oxide layer during planarization.

Brief Description of the Drawing Figures

The present invention will hereinafter be described in conjunction with the appended drawing figures, wherein like numerals denote like elements, and:

30   Figure 1 is a perspective view of an exemplary CMP machine upon which the probe of the present invention is suitably installed;

3

Figure 2 is a top view of the CMP machine of Figure 1, showing an exemplary orientation of the probe assemblies of the present invention;

Figure 3 is a schematic view of a probe in accordance with the present invention configured to sample the surface of a wafer, including a light source, photospectrum meter,
5   controller, and display;

Figure 4 is an end view of an exemplary bifurcated probe assembly having a plurality of illuminators and a receiver probe;

Figure 5 is a schematic view of the probe of **Figures 3 and 4** sampling an individual die structure of a wafer;

10   Figure 6 is a view of one side of a wafer having a plurality of microelectronic die structures disposed thereon;

Figure 7 is a top view of an exemplary embodiment of a wafer carrier lock-in mechanism in accordance with the present invention;

Figure 8 is a side view of the lock-in mechanism of **Figure 7**;

15   Figure 9 is a side view of the lock-in mechanism of **Figures 7 and 8** with a carrier and wafer assembly in operative engagement with the lock-in mechanism;

Figure 10 is a cross sectional depiction of an exemplary semiconductor wafer;

Figure 11 is a schematic view of a probe sampling a metallic layer of the semiconductor wafer shown in **Figure 10**;

20   Figure 12 shows an exemplary output signal generated in response to the sampling of the metallic layer depicted in **Figure 11**;

Figure 13 is a schematic view of a probe sampling an oxide layer of the semiconductor wafer shown in **Figure 10**; and

Figure 14 shows an exemplary output signal generated in response to the sampling
25   of the oxide layer depicted in **Figure 13**.

Detailed Description of Preferred Exemplary Embodiments

The subject invention relates to the in-process detection of characteristics of a layer on a workpiece using a broad spectrum light source, a photospectrum meter, and a controller including a smart algorithm for translating the output of the photospectrum meter to a human
30   readable display relating to the workpiece being examined. The preferred embodiment set forth herein relates to the detection of oxide layer thickness on a semiconductor wafer (either patterned or non-patterned); it will be appreciated, however, that the principles of the present

invention may be employed to ascertain any number of characteristics associated with a workpiece surface, including end point detection, the detection of surface irregularities, planarity, and the like.

Referring now to **Figures 1-2**, a wafer polishing apparatus 100 is shown embodying
5 the present invention. Wafer polishing apparatus 100 suitably comprises a multiple head wafer polishing machine which accepts wafers from a previous processing step, polishes and rinses the wafers, and reloads the wafers back into wafer cassettes for subsequent processing.

Discussing now the polishing apparatus 100 in more detail, apparatus 100 comprises an unload station 102, a wafer transition station 104, a polishing station 106, and a wafer rinse
10 and load station 108.

In accordance with a preferred embodiment of the invention, cassettes 110, each holding a plurality of wafers, are loaded into the machine at unload station 102. Next, a robotic wafer carrier arm 112 removes the wafers from cassettes 110 and places them, one at a time, on a first wafer transfer arm 114. Wafer transfer arm 114 then sequentially lifts and moves each
15 wafer into wafer transition section 104. That is, transfer arm 114 suitably places an individual wafer on one of a plurality of wafer pick-up stations 116 which reside on a rotatable table 120 within wafer transition section 104. Rotatable table 120 also suitably includes a plurality of wafer drop-off stations 118 which alternate with pick-up stations 116. After a wafer is deposited on one of the plurality of pick-up stations 116, table 120 rotates so that a new station 116 aligns
20 with transfer arm 114. Transfer arm 114 then places the next wafer on the new empty pick-up station 116. This process continues until all pick-up stations 116 are filled with wafers. In the illustrated embodiment of the invention, table 120 includes five pick-up stations 116 and five drop-off stations 118.

Next, a wafer carrier apparatus 122, comprising individual wafer carrier elements
25 124, suitably aligns itself over table 120 so that respective carrier elements 124 are positioned directly above the wafers which reside in respective pick-up stations 116. The carrier apparatus 122 then drops down and picks up the wafers from their respective stations and moves the wafers laterally such that the wafers are positioned above polishing station 106. Once above polishing station 106, carrier apparatus 122 suitably lowers the wafers, which are held by
30 individual elements 124, into operative engagement with a polishing pad 126 which sits atop a lap wheel 128. During operation, lap wheel 128 causes polishing pad 126 to rotate about its vertical axis. At the same time, individual carrier elements 124 spin the wafers about their

respective vertical axes and oscillate the wafers back and forth across pad 126 (substantially along arrow 133) as they press against the polishing pad. In this manner, the under surface of the wafer is polished or planarized.

After an appropriate period of time, the wafers are removed from polishing pad 126,

5   and carrier apparatus 122 transports the wafers back to transition station 104. Carrier apparatus 122 then lowers individual carrier elements 124 and deposits the wafers onto drop-off stations 118. The wafers are then removed from drop-off stations 118 by a second transfer arm 130. Transfer arm 130 suitably lifts each wafer out of transition station 104 and transfers them into wafer rinse and load station 108. In the load station 108, transfer arm 130 holds the wafers

10   while they are rinsed. After a thorough rinsing, the wafers are reloaded into cassettes 132, which then transports the subsequent stations for further processing or packaging.

Although CMP machine 100 is shown having five polishing stations, it will be appreciated that the present invention may be employed in the context of virtually any number of polishing stations. Moreover, the present invention may also be employed in circumstances

15   where not all of the polishing stations are functioning at a time. For example, many standard wafer cassettes carry twenty-four individual workpieces in a single cassette. Consequently, because there are often five workpiece chucks on a single CMP machine, often times the last four disks within a cassette are polished at one time, leaving the fifth disk-holder empty.

With continued reference to Figure 2, a probe assembly 127 is suitably configured

20   near the outer perimeter edge of polishing pad 126 proximate each carrier element 124. More particularly, in a preferred embodiment of the present invention, each respective carrier element suitably oscillates back and forth along arrow 133; each carrier element 124 also suitably rotates a workpiece about the vertical axis of carrier element 124. At the same time, lap wheel 128 and pad 126 are advantageously configured to rotate about their vertical axis, for example, in a

25   counter clockwise direction as indicated by arrow 134.

In accordance with a particularly preferred embodiment, each carrier element 124 is suitably configured to periodically extend radially outward from the center of table 126 along arrow 125 such that at least a portion of the outside radius of each workpiece extends beyond the outer edge 137 of table 126. By crossing the outer edge of polishing table 126, surface

30   material thicknesses, desirable material removal rates, and the extent of layer removal may be obtained for the workpieces. As a workpiece extends beyond the outer perimeter of the polishing pad, along line 125, the bottom facing surface of the workpiece may be conveniently

optically engaged by probe assembly 127, as described in greater detail below in conjunction with **Figure 3**.

In accordance with a further aspect of the present invention, apparatus 100 may be configured with a probe assembly 129 useful for detecting the presence of a wafer or wafer fragment on polishing pad 126 during the polishing process. In accordance with this aspect of the invention, if a wafer or wafer fragment is detected on the pad at an inappropriate time, the CMP machine 100 will shut down. A detailed discussion of the operation of probe assembly 129 is discussed in detail in Holzapfel et al., U.S. Patent Application Serial No. 08/683,150, filed on July 17, 1996, and entitled *Methods and Apparatus for the In-Process Detection of Workpieces in a CMP Environment*, which is incorporated herein by reference.

Referring now to **Figure 3**, an exemplary embodiment of probe assembly 127, in accordance with the present invention, suitably comprises a housing 310 having a nozzle 312 through which compressed air is suitably directed at the under surface of a workpiece (*e.g.*, semiconductor wafer disk) 306, a bifurcated fiber probe 316, a light source 322, a photospectrum analyzer 324, a controller/processor 326, and a display 328.

As best seen in **Figure 3**, an exemplary workpiece 306 is shown being polished by polishing pad 126 as described above in conjunction with **Figures 1 and 2**. For clarity, carrier element 124 and other components are omitted from **Figure 3**. Probe assembly 127 is suitably mounted proximate the outer perimeter 330 of pad 126, such that nozzle 312 may be directed at a bottom surface 304 of workpiece 306 when at least a portion of the workpiece extends off the perimeter edge of polishing pad 126.

In accordance with a particularly preferred embodiment of the present invention, a supply of compressed air, for example in the range of 0 to 20 PSI and most preferably about 5 PSI, is urged through housing 310 and nozzle 312 to clear away slurry from an exemplary region 314 on undersurface 304 under examination. As the compressed air clears away slurry from the underside of the workpiece, probe 316 suitably outputs a broad band (*e.g.*, white light) light source at region 314; a portion of the light emitted by probe 316 is reflected or scattered back from region 314 and captured by probe 316. In a preferred embodiment, the light output by probe 316 suitably passes through a collimating lens 408 which collimates the light (*see* **Figure 5**). Thus, while a cross-sectional area of light source (probe) 316 is suitably in the range of 0.1 to 10.0 square millimeters and preferably about 1.0 square millimeter, collimating lens 408 is suitably configured to project the light to cover a region 314 on the-wafer undersurface

that suitably comprises an area in the range of about 10.0 to about 30.0 square millimeters, and preferably about 20.0 square millimeters.

More particularly and with continued reference to **Figures 3 and 4**, probe 316 suitably comprises a plurality of light illuminators 350 and a single receiver probe 352. In

5   accordance with this aspect of the present invention, probe 316 preferably comprises a plurality of (*e.g.* six) illuminators 350 suitably configured around a single receiver probe 352 disposed in the center of the illuminators. In accordance with this aspect of the present invention, illuminators 350 may be suitably be grouped in a hexagonal configuration. The diameter of each illuminator 350 and the receiver probe 352, as shown in **Figure 4**, is suitably about 100 to about

10   300 microns and preferably about 200 microns. Accordingly, the diameter of probe 316 is suitably in the range of .1 to 5 millimeters, and preferably about .5 to about 2 millimeters, and most preferably about 1 millimeter. Probe 316 further comprises a transmitter cable 318 through which light is transmitted from light source 322 to illuminators 350 of probe 316 and onto the undersurface of the workpiece. Similarly, probe 316 suitably comprises a receptor

15   cable 320 which receives light from receiver probe 350 and transmits it to photospectrum meter 324. It will be appreciated that the undersurface of the workpiece may be sampled by probe assembly 127 at any desired rate or the sampling may be substantially continuous.

Although the preferred embodiment of light source 322 has been described in accordance with **Figures 3 and 4**, light source 322 may suitably comprise any source capable

20   of applying a desired light signal (*e.g.* broadband, narrow band, or substantially monochromatic) to the surface of the workpiece. For example, any suitable source (*e.g.* a tungsten halogen light source) capable of omitting a broad band light signal, for example in the range of 350 to 2000 nanometers, and most preferably in the range of 400 to 850 nanometers, is acceptable. In accordance with the present invention, a suitable halogen light source may comprise a model

25   number L73A98, available from the Gilway Corporation of Massachusetts.

Although cables 318 and 320 suitably comprise fiberoptic cables in the preferred embodiment, virtually any conductor may be employed which satisfactorily delivers an appropriate signal (*e.g.* a light signal) to the workpiece and captures at least a portion of the signal reflected by the workpiece. Moreover, although the preferred embodiment set forth

30   herein employs a light signal, virtually any convenient modality may be employed to interrogate the surface of the workpiece, *e.g.*, an acoustic signal, magnetic signal, or the like.

Photospectrum meter 324 suitably comprises any circuit capable of interpreting the

signal reflected from the undersurface of the workpiece. In a preferred embodiment, photospectrum meter 324 suitably comprises a PCMCIA-based photospectrum meter model number PS1000 available from the Mission Peak Optics Company of Fremont, California.

Controller 326 suitably comprises any general purpose controller capable of
5   receiving an output signal 332 from meter 324 and calculating various parameters from signal 332. In the preferred embodiment, controller 326 is suitably configured to interpret signal 332 and thereby derive the thickness of the oxide layer present in region 314 of workpiece 306. In a particularly preferred embodiment, controller 326 suitably comprises any general purpose personal computer, for example a PC, available from the Mission Peaks Optics Company of
10 Fremont, California.

Controller 326 is also suitably configured to output a signal 334 to display terminal 328. In a preferred embodiment, signal 334 is indicative of the thickness of the oxide layer at region 314; it will be understood, however, that signal 334 may embody any suitable information or characteristics about surface 304 of the workpiece, such that any number of
15 parameters may be conveniently displayed on the screen associated with display module 328. For example, signal 344 may convey information indicative of the removal of a first semiconductor layer from a second semiconductor layer (described below). In the embodiment illustrated in Figure 3, a graph of oxide layer thickness versus time is shown.

With continued reference to Figure 3 display terminal 328 may be suitably
20 configured to display information pertaining to the undersurface of the workpiece. (e.g., the thickness of the oxide layer or the material composition of the exposed layer) in any desired format. In the thickness versus time graph shown in Figure 3, the remaining processing time necessary to arrive at a desired thickness 340 may be visually assessed by the operator; alternatively, controller 326 may be configured to "predict" the time necessary to arrive at a
25 desired thickness for a given pressure and also to display the remaining time to the operator. Alternatively, the controller may be configured to transmit a second output signal 342 to the main controller of machine 100, for example to vary the pressure or other operating parameter(s) associated with the particular carrier element 124 corresponding to the workpiece under inspection. For example, if it is desired that all workpieces complete their processing at
30 approximately the same time, and wherein one or more of the workpieces are closer to the desired thickness than other workpieces, it may be advantageous to reduce the pressure for those

workpieces where less material remains to be removed and/or to increase the pressure for those workpieces where a relatively larger amount of material remains to be removed.

In this regard, the present inventors have determined that typical desired material removal rates of oxide layers on semiconductor wafers generally range from 1,000-5,000

5   angstroms per minute, and preferably about 2,500 angstroms per minute. By calculating the differences in thickness over different sampling periods, controller 326 may also be suitably configured to generate a real time or average material removal rate. In accordance with a preferred embodiment, controller 326 may suitably be configured to output signal 342 to increase or decrease the removal rate, as desired.

10   With continued reference to **Figure 3**, probe assembly 127 may be mounted to machine 100 in any convenient way, for example, by attaching probe assembly 127 to the frame associated with machine 100 by any suitable fastening mechanism. Indeed, it may be possible to dispose respective probe housings 310 quite close to the surface of the workpiece, for example in the range of 0.1 to 0.5 inches and most preferably about .3 inches from the

15   workpiece. Even though this environment may be sprayed by slurry droplets from time to time, the compressed air ejected from housing 310 by nozzle 312 suitably substantially prevents slurry from entering the housing and corrupting probe 316. One preferred embodiment of an exemplary mounting mechanism is discussed in more detail below in conjunction with **Figures 7-9**.

20   In accordance with a particularly preferred embodiment, probe assembly 127 may be suitably configured to output signal 342 to machine 100 to thereby terminate the processing of a particular workpiece when it is determined that desired thickness 340 has been achieved. In this way, although it still may be desirable to verify the thickness of the oxide layer once the workpieces have been removed, a very high degree of accuracy in the actual thickness of the

25   oxide layer is obtained. In accordance with this aspect of the present invention, the need to place partially completed disks back onto machine 100 for further material removal is substantially eliminated. Similarly, the risk of removing too much of the oxide layer, thus degrading the wafers, is also greatly reduced. In this respect, the present invention may be alternatively configured to indicate when the oxide layer is initially exposed (described below).

30   The manner in which probe assembly 127 samples and interprets the scattered light signals to determine wafer surface thickness will now be described in conjunction with **Figures 5 and 6**.

Referring now to **Figure 6**, an exemplary embodiment of a wafer surface comprises a plurality of substantially similar die structures arranged in a rectangular grid pattern. As shown in **Figure 5**, each individual die structure 406 may comprise in schematic cross section, one or more alternating substrate and oxide layers; for example, a substrate layer 404 and an

5 oxide layer 402. Substrate layer 404 generally comprises a plurality of microelectronic structures substantially defining a substrate topology 405. Because the surface layer or topology 405 of substrate 404 is non-uniform, it is very difficult to accurately determine the thickness of oxide layer 402 at any particular point on the surface of wafer 400. Therefore, to obtain accurate oxide thickness readings, the effect of the non-uniform substrate surface must be

10 minimized or otherwise accounted for. An exemplary wafer surface sampling and analysis method in accordance with the present invention will now be discussed in greater detail.

In accordance with a particularly preferred embodiment of the present invention, light is transmitted from illuminators 350 of probe 316 through collimating lens 408 and onto the undersurface of wafer 400. Generally, part of the transmitted light will be reflected or

15 scattered from oxide layer 402 back to receiver probe 352. However, a substantial portion of the light passes through oxide layer 402 and reflects off the substrate layer 404 (and more particularly, the non-uniform surface 405). As the light is reflected off the substrate and oxide layers, it passes back through collimating lens 408, which essentially focuses the reflected light back to receiver probe 352. The reflected light then passes to photospectrum meter 324 through

20 fiberoptic cable 320. Photospectrum meter 324 then divides the light into discrete bands of predetermined frequency (or wavelength) ranges and converts the light frequency signals into a digital output signal 332 which is communicated to processor/controller 326.

In accordance with a preferred embodiment, processor 326 analyses the converted light frequency signals according to the well-known photo-interference technique using Fresnel's

25 Equation to obtain the thickness of the oxide layer at the sampled area of the wafer surface. In accordance with this aspect of the present invention, the relationship:

$$\frac{4\pi n d}{\lambda}$$

may be conveniently solved to obtain the oxide thickness, where d = the oxide thickness; n = the refractive index of the sample material (e.g. the oxide layer); and λ = the

30 wavelength of the light.

However, as discussed previously, because the topology of substrate surface 405 is non-uniform, it is very difficult to get an accurate measurement of the thickness of oxide layer

402. Thus, an averaging technique is desirably employed which effectively cancels out many of the effects attributable to the complexity of the substrate layer topology.

The averaging technique will now be discussed in greater detail in conjunction with Figures 5 and 6. Referring now to Figure 6, as discussed previously, an exemplary

5 embodiment of wafer 400 comprises a plurality of die structures 406, each comprising substantially similar or identical substrate topologies. Because of the repeating nature of the die structures, the wafer surface may be advantageously sampled for approximately one full wafer rotation, and the measurements taken during that rotation suitably averaged to largely suppress or even cancel out the effects of the non-uniform topology of the dies. That is, while the

10 sampling of only one die may yield an inaccurate reading of the thickness of the oxide layer, by averaging the sample readings obtained from many similar die structures for one full wafer rotation, the non-uniformity of each die will be effectively canceled out by the averaging technique, thus giving a more accurate oxide thickness reading.

In accordance with this aspect of the present invention, probe 316 suitably collects

15 between about 100 and about 300 samples for one complete rotation, and preferably about 200 samples. Further, one complete rotation of the wafer generally takes approximately 2 seconds; thus, the sampling rate of probe 316 is suitably between about 100 to about 300 samples per second and preferably about 200 samples per second.

The data for each sample is transmitted to processor 326 which stores and

20 accumulates the data. After sampling a portion (e.g. region 314) of the wafer for approximately one full wafer rotation, the accumulated data (approximately 200 samples) is averaged and one average oxide thickness is calculated, for example using the aforementioned Fresnel technique. In accordance with a particularly preferred embodiment of the invention, all the sampled data for one rotation is added together and averaged and then the oxide layer thickness is calculated

25 from the averaged data. Alternatively, in accordance with yet another preferred embodiment of the invention, an oxide thickness may be calculated for each individual data sample, and thereafter an average thickness calculated from all the individually calculated thicknesses.

In accordance with a particularly preferred embodiment, probe assembly 127, and in particular probe 316, is suitably configured to sample the wafer in a substantially circular path

30 410 (see Figure 6) so that each die 406 along path 410 is sampled in a substantially uniform manner. That is, because the die structures on the wafer are in a substantially uniform grid pattern and because the field of view of probe 316 suitably corresponds to approximately one

die, the probe is likely to sample a complete die structure (as opposed to simply sampling sections of multiple die structures) as it traverses the wafer in a substantially circular pattern. In accordance with this aspect of the present invention, a wafer carrier lock-in mechanism (discussed in detail below) may be used to ensure that the field of view of the probe follows a

5   substantially circular path around the wafer, thus eliminating a spiral reading effect.

Wafer carrier lock-in mechanism 500 will now be discussed in greater detail in conjunction with Figures 7-9. An exemplary embodiment of wafer lock-in mechanism 500 preferably comprises a base 502, a rotatable carrier guide 504, a spring 506 and a stopper 508. Rotatable carrier guide 504 is suitably mounted to base 502 with a bearing assembly 510 to

10   permit free rotation of guide 504. Further, spring 506 is suitably mounted between base 502 and stopper 508. Finally, probe 316 is securely mounted within base 502 so that the illuminators and receptor probe are pointed upward toward wafer 400 and wafer carrier element 124.

During operation, wafer carrier 124 rotates about its vertical axis and oscillates back

15   and forth across polishing pad 126. As carrier element 124 oscillates across the pad, a portion of the carrier element periodically extends beyond the edge of the pad, contacting rotating carrier guide 504. Carrier guide 504 suitably rotates about bearing 510 as carrier element 124 rotates, thus minimizing friction between the two elements. Once carrier element 124 contacts guide 504, the field of view of probe 316 becomes fixed at a specific radial point on wafer 400.

20   In accordance with this aspect of the invention, as the carrier and wafer assembly rotate, the field of view of the probe traverses a substantially circular path around the wafer ensuring relatively accurate readings.

In accordance with an exemplary embodiment of the present invention, as the carrier and wafer assembly continue to oscillate further out from the pad, the carrier pushes probe 316,

25   guide 504, and base 502 assembly towards stopper 508, compressing spring 506. Then, as the carrier and wafer assembly begin to oscillate back towards the center of polishing pad 126, the tension in spring 506 causes base 502, carrier guide 504, and probe 316 assembly to remain in contact with and to move with the carrier element and wafer, thus maintaining the position of the probe's field of view on the wafer. Accordingly, as the lock-in mechanism 500 moves with

30   the carrier and wafer assembly, the probe maintains a substantially circular field of view around the wafer as the carrier and wafer rotate and oscillate back and forth across the pad, and

13

therefore, preventing the probe from sampling along a less desirable spiral path on the surface of the wafer.

As mentioned above, the present invention may also be employed to detect when a first material layer, e.g., a tungsten or titanium layer, has been removed from a second material

5   layer, e.g., an oxide layer. It should be noted that the present invention may be suitably adapted to detect semiconductor, conductor, or other layers that may be present on a semiconductor wafer. Semiconductor layers are described herein for illustrative purposes only, and the present invention is not limited to the detection of such layers. **Figure 10** is a cross sectional representation of an exemplary semiconductor wafer 600. Of course, the dimensions of wafer

10   600 are exaggerated for illustrative purposes. Wafer 600 may include a silicon substrate base 602, an oxide (silicon dioxide) layer 604, a titanium layer 606, a titanium nitride layer 608, and a tungsten layer 610. In accordance with conventional semiconductor fabrication techniques, a number of plugs 611 may be formed within wafer 600.

During planarization, the metallic upper layers, e.g., tungsten layer 610, titanium

15   nitride layer 608, and titanium layer 606, are removed from wafer 600, while some material is maintained within the associated plugs 611. Preferably, material removal terminates (or slows down) when oxide layer 604 is exposed. As described above, the present invention may be employed to measure the thickness of oxide layer 604 to optimize the planarization process. Those skilled in the art will appreciate that the specific layering configuration and the

20   composition of the various layers may vary from wafer to wafer depending on the device being created.

For purposes of the following description, a first region 612 of wafer 600 may be defined as that portion of wafer 600 above oxide layer 604. A second region 614 may be defined as that portion of wafer 600 that includes oxide layer 604, silicon base 602, and any

25   other layers (not shown) that may be formed below oxide layer 604. It should be appreciated that although the transition from titanium layer 606 to oxide layer 604 is described herein, the present invention may be suitably adapted to detect the transition from any two material layers having distinguishable optical reflective characteristics. Such adaptation may require additional or alternative processing and/or the application of different detection algorithms than those

30   described above.

The metallic layer endpoint detection feature takes advantage of the distinguishable reflective characteristics of the various semiconductor layers within wafer 600. **Figure 11**

shows probe 316 (see Figures 3 and 5) directing an input interrogation signal 616 at wafer 600. Input signal 616 may be directed toward wafer 600 at an angle of approximately 80 to 100 degrees, and preferably at an angle of approximately 90 degrees, relative to the upper surface 620 of wafer 600. The "metallic" layer 618 shown in Figure 11 is intended to generically

5   indicate any semiconductor material of a metallic composition, e.g., tungsten, copper, titanium, titanium nitride, or the like. The metallic nature of layer 618 causes input signal 616 to reflect from layer 618.

Figure 12 depicts an exemplary output 622 that may be generated in response to the continued presence of metallic layer 618 and displayed on display terminal 328 (see Figure 3

10 and accompanying discussion). A flat output (or any other predetermined display or indicator) may indicate that the reflected signal detected at probe 316 is less than a predetermined threshold value. Alternatively, the same output may indicate that the reflected signal detected at probe 316 has certain detectable characteristics that are distinguishable from a signal reflected from oxide layer 604. For example, the reflected signal itself may have monotonic characteristics when

15 reflected from metallic layer 618 and sinusoidal characteristics when reflected from oxide layer 604 (or a nonmetallic layer).

In accordance with a further aspect of the present embodiment of the invention, in addition to generating such a display, wafer polishing apparatus 100 may be instructed, by suitable processing and control elements, to continue with the planarization procedure, i.e., to

20 continue removing additional layers or more material from the present layer.

Figure 13 shows probe 316 directing input signal 616 at wafer 600 after the planarization process has exposed oxide layer 604. As shown, the metallic layer 618 has been removed from wafer 600. Although not shown in Figure 13, various plugs 611 formed from tungsten may remain embedded in oxide layer 604 (see, for example, Figure 10). Figure 14

25 depicts an exemplary output 624 that may be displayed on display terminal 328 after oxide layer 604 has become exposed. A sinusoidal output (or any other predetermined display or indicator) may indicate that the reflected signal detected at probe 316 is greater than the predetermined threshold value. Alternatively, the same output may indicate that the reflected signal detected at probe 316 has certain detectable characteristics that are distinguishable from a signal reflected

30 from metallic layer 618. For example, the reflected signal may become sinusoidal in nature after the metallic layer 618 has been cleared away and oxide layer 604 becomes the reflective surface. In addition, suitable control processes may cause wafer polishing apparatus 100 to halt

the planarization procedure or slow the procedure down such that subsequent removal of material may be closely monitored.

It should be appreciated that the detection of oxide layer 604 in this manner is substantially independent of the amount and type of slurry and other consumables that may be
5   present in the CMP environment. Unlike conventional endpoint detection techniques that depend on mechanical characteristics such as polishing pad friction and motor currents, the present invention can effectively detect the removal of a metallic layer without relying upon physical interaction with the system. Those skilled in the art will appreciate that the present invention may be utilized to detect the transition between any two reflectively distinguishable materials
10   during a removal or planarization process. For example, if required, the processing and display functions of the present invention may be appropriately modified to detect the removal of an oxide layer from a metallic layer, or one metallic layer from another metallic layer.

In summary, the present invention provides improved methods and apparatus for measuring the thickness of layers on semiconductor wafers and other workpiece surfaces. The
15   present invention facilitates the in-process, in-situ, substantially real time measurement of the actual thickness of a surface layer of a workpiece under inspection. A smart algorithm is employed to calculate the thickness of the oxide layer from information gathered from light signals reflected from the surface of the wafer. In a preferred embodiment, the oxide layer thickness as a function of time is displayed on a view screen for convenient observation by the
20   operator of the machine. In accordance with a different embodiment of the invention, the present invention may be alternatively configured to optically detect the endpoint when a metallic layer is removed from the oxide layer during the planarization process.

Although the subject invention is described herein in conjunction with the appended drawing figures, it will be appreciated that the invention is not limited to the specific form
25   shown. Various modifications in the selection and arrangement of parts, components, and processing steps may be made in the implementation of the invention. For example, although a preferred embodiment is set forth in which a tungsten halogen light source is used in connection with fiberoptic conductors, it will be appreciated that virtually any interrogation signal may be employed through appropriate conductors, such that in-process, in-situ monitoring
30   of workpiece surface parameters are made available for analyses. Moreover, although the light source, photospectrum meter, controller, and terminal display are illustrated in Figure 3 in schematic form, it will be appreciated that only the probe 316 need be disposed proximate the

16

workpieces various of the other components, including the light source, photospectrum meter, controller, and screen display may be disposed remotely from the workpiece, as desired. These and other modifications may be made in the design and arrangement of the various components which implement the invention without departing from the spirit and scope of the invention as

5   set forth in the appended claims.

CLAIMS:

1.   An apparatus for measuring the thickness of an oxide layer on a semiconductor wafer during CMP processing of said wafer, comprising:

        a probe assembly disposed proximate a polishing pad associated with a CMP machine, such that said probe establishes intermittent communication with said wafer as said wafer is processed on said polishing pad;

        a signal source associated with said probe assembly, configured to direct an interrogation signal at a location on a surface of said wafer while said wafer is rotating, wherein said location may be any location on the surface of said wafer;

        a receptor configured to capture a portion of said interrogation signal which is reflected from the surface of said wafer; and

        a processor configured to process said captured interrogation signal and generate an output indicative of a parameter associated with said wafer;

        wherein said apparatus samples a plurality of said locations on the surface of said wafer using a plurality of said interrogation signals, and wherein said output is an average output calculated by said processor from said plurality of said interrogation signals.

2.   The apparatus of claim 1, wherein said parameter corresponds to a thickness of an oxide layer on the surface of said wafer.

3.   The apparatus of claim 1, wherein said signal source comprises a white light source.

4.   The apparatus of claim 3, wherein said processor comprises a photospectrum meter for processing said captured light signal, and a controller for calculating said parameter based on the output of said photospectrum meter.

5.   The apparatus of claim 1, further comprising a display terminal for displaying a human readable output indicative of said parameter associated with said wafer.

6.   The apparatus of claim 1, further comprising a nozzle disposed proximate said probe and configured to apply compressed air to the surface of said wafer in the vicinity of said probe.

7.   A method for measuring the thickness of a layer on the surface of a workpiece in a CMP process, comprising the steps of:

(a)   rotating and moving said workpiece on a surface of a rotating polishing table such that a portion of said workpiece is caused to intermittently extend beyond an outer perimeter of said polishing table;

(b)   disposing a probe proximate said outer perimeter of said polishing table such that said probe engages said workpiece in a noncontacting manner when said portion of said workpiece extends beyond said perimeter of said polishing table;

(c)   applying an interrogation signal from said probe to each one of a plurality of locations on said portion of said workpiece while said workpiece is rotating, so that said plurality of locations on said portion of said workpiece are sampled, wherein said plurality of locations may be any location on said portion of said workpiece;

(d)   capturing a portion of said interrogation signals reflected from each of said plurality of locations on said workpiece; and

(e)   analyzing said reflected interrogation signals from each of said plurality of locations on said workpiece and calculating a real time output parameter pertaining to said workpiece, wherein said output parameter is an average output parameter calculated from said interrogation signals.

8.   The method according to claim 7, wherein said probe is configured to optically engage said workpiece.

9.   The apparatus of claim 1, wherein said apparatus samples said plurality of locations on the surface of said wafer for one complete revolution of said wafer.

10.   The apparatus of claim 9, wherein said apparatus samples between 100 and 300 locations on the surface of said wafer during said one complete revolution of said wafer.

11.   The apparatus of claim 10, wherein said sampling rate of said probe is about 50 to 150 samples per second.

12.   The apparatus of claim 1, wherein said location is a patterned die structure on said wafer.

13.   The apparatus of claim 2, wherein the thickness of the oxide layer is an average thickness of the oxide layer on the surface of said wafer, and is calculated by first averaging said interrogation signals to create an average interrogation signal value and then calculating the average thickness of the oxide layer from the average interrogation signal value.

14.   The apparatus of claim 2, wherein the thickness of the oxide layer is an average thickness of the oxide layer on the surface of said wafer, and is calculated by first

19

calculating an oxide layer thickness value for each of said interrogation signals and then averaging the calculated oxide layer thickness values to derive an average thickness of the oxide layer.

15.   The apparatus of claim 3, wherein said light source is a white light source.

16.   The apparatus of claim 15, wherein said light source is a tungsten halogen light source.

17.   The apparatus of claim 1, wherein said output from said processor is transmitted to a CMP system processor which uses said output in a feedback loop to control wafer process parameters.

18.   The apparatus of claim 17, wherein said CMP system processor uses said output to calculate an average material removal rate for said wafer.

19.   The apparatus of claim 18, wherein said CMP system processor calculates an estimated time to finish processing of said wafer based on said average material removal rate.

20.   The apparatus of claim 18, wherein said CMP system processor causes at least one processing parameter to vary, thereby changing said average material removal rate.

21.   The apparatus of claim 18, wherein said CMP system processor causes a wafer rotation rate to vary, thereby changing said average material removal rate.

22.   The apparatus of claim 17, wherein said output corresponds to an oxide layer thickness on the surface of said wafer, and said CMP system processor is configured to stop processing of said wafer when said oxide layer reaches a desired thickness.

23.   The method of claim 7, wherein step (c) is performed for one complete revolution of said workpiece.

24.   The method of claim 23, wherein the surface of said workpiece is sampled at between 100 and 300 locations during said one complete revolution of said workpiece.

25.   The method of claim 7, wherein said output parameter corresponds to a thickness of an oxide layer on the surface of said workpiece, and wherein step (e) further comprises the steps of :

averaging said interrogation signals from each of said plurality of locations on said workpiece to create an average interrogation signal value; and

calculating an average thickness of the oxide layer from the average interrogation signal value.

26.   The method of claim 7, wherein said output parameter corresponds to a thickness of an oxide layer on the surface of said workpiece, and wherein step (e) further comprises the steps of :

calculating an oxide layer thickness value for each of said interrogation

5   signals; and

averaging the calculated oxide layer thickness values to derive an average thickness of the oxide layer.

27.   An apparatus for measuring the thickness of an oxide layer on a semiconductor wafer during CMP processing of said wafer, comprising:

10   a probe assembly disposed proximate a polishing pad associated with a CMP machine, wherein said probe assembly establishes intermittent communication with said wafer as said wafer is processed on said polishing pad, said probe assembly comprising:

a cluster of a plurality of signal sources configured to direct an

15   interrogation signal at a location on a surface of said wafer while said wafer is rotating, wherein said location may be any location on the surface of said wafer; and

a receptor located at the center of said cluster of said plurality of signal sources and configured to capture a portion of said interrogation

20   signal which is reflected from the surface of said wafer; and

a processor configured to process said captured interrogation signal and generate an output indicative of a parameter associated with said wafer;

wherein said apparatus samples a plurality of said locations on the surface of said wafer using a plurality of said interrogation signals from said probe, and

25   wherein said output is an average output calculated by said processor from said plurality of said interrogation signals.

28.   The apparatus of claim 27, wherein six of said signal sources are disposed around said receptor in a hexagon configuration.

29.   In a planarization system configured to process a semiconductor wafer

30   initially having at least a first material layer formed over a second material layer, a method for detecting the removal of said first material layer from said semiconductor wafer during operation of said planarization system, said method comprising the steps of: _ _ _

directing an input signal at a portion of said semiconductor wafer;

detecting a reflected signal produced in response to said input signal;

analyzing characteristics of said reflected signal; and

indicating removal of said first material layer from said semiconductor wafer in

5   response to characteristics of said reflected signal.

30.   A method according to claim 29, wherein said indicating step is performed
when said reflected signal exhibits monotonic characteristics.

31.   A method according to claim 29, wherein said indicating step is performed
when said reflected signal has sinusoidal characteristics.

10   32.   A method according to claim 29, further comprising the step of displaying
a human readable output indicative of removal of said first material layer from said
semiconductor wafer.

33.   A method according to claim 29, further comprising the step of adjusting
an operational parameter of said planarization system in response to said indicating step.

15   34.   An apparatus for detecting removal of a first material layer initially
formed over a second material layer on a semiconductor wafer during a planarization procedure,
said apparatus comprising:

means for generating an input signal for establishing optical communication with a
surface of said semiconductor wafer during said planarization procedure;

20   a receptor configured to receive a reflected signal produced in response to said input
signal, said reflected signal being reflected from said surface of said semiconductor wafer; and

a processor configured to process said reflected signal and to produce a first output
signal having a first characteristic indicative of the presence of said first material layer and a
second output signal having a second characteristic indicative of the presence of said second

25   material layer.

35.   An apparatus according to claim 34, wherein said means for generating
comprises a substantially white light source.

36.   An apparatus according to claim 34, wherein said processor produces said
first output signal when said reflected signal is less than a predetermined threshold value.

30   37.   An apparatus according to claim 34, wherein said processor produces said
second output signal when said reflected signal exhibits monotonic characteristics.

38.   An apparatus according to claim 34, wherein said processor produces said second output signal when said reflected signal has sinusoidal characteristics.

39.   An apparatus according to claim 34, further comprising a display terminal for displaying a human readable output indicative of removal of said first material layer from said semiconductor wafer.

40.   An apparatus according to claim 34, wherein said semiconductor wafer has a relatively planar upper surface and said means for generating is positioned such that said input signal is directed toward said semiconductor wafer at an angle between 80 and 110 degrees relative to said upper surface.

41.   An apparatus according to claim 34, wherein said first material layer comprises a material selected from the group consisting of copper, tungsten, titanium, and titanium nitride.

42.   An apparatus according to claim 34, wherein said second material layer comprises an oxide composition.

43.   A method for detecting removal of a first material layer initially formed above a second material layer on a semiconductor wafer, said method comprising the steps of:

applying an interrogation signal to an exposed portion of said first material layer;

conducting a planarization procedure on said first material layer;

detecting a signal reflected from said second material layer and produced in response to said interrogation signal, said second material layer having different optical interference characteristics relative to said first material layer; and

indicating removal of said first material layer from said semiconductor wafer in response to said signal.

44.   A method according to claim 43, further comprising the step of halting said planarization procedure in response to said indicating step.

45.   A method according to claim 43, wherein said first material layer comprises a metallic composition and said second material layer comprises an oxide composition.
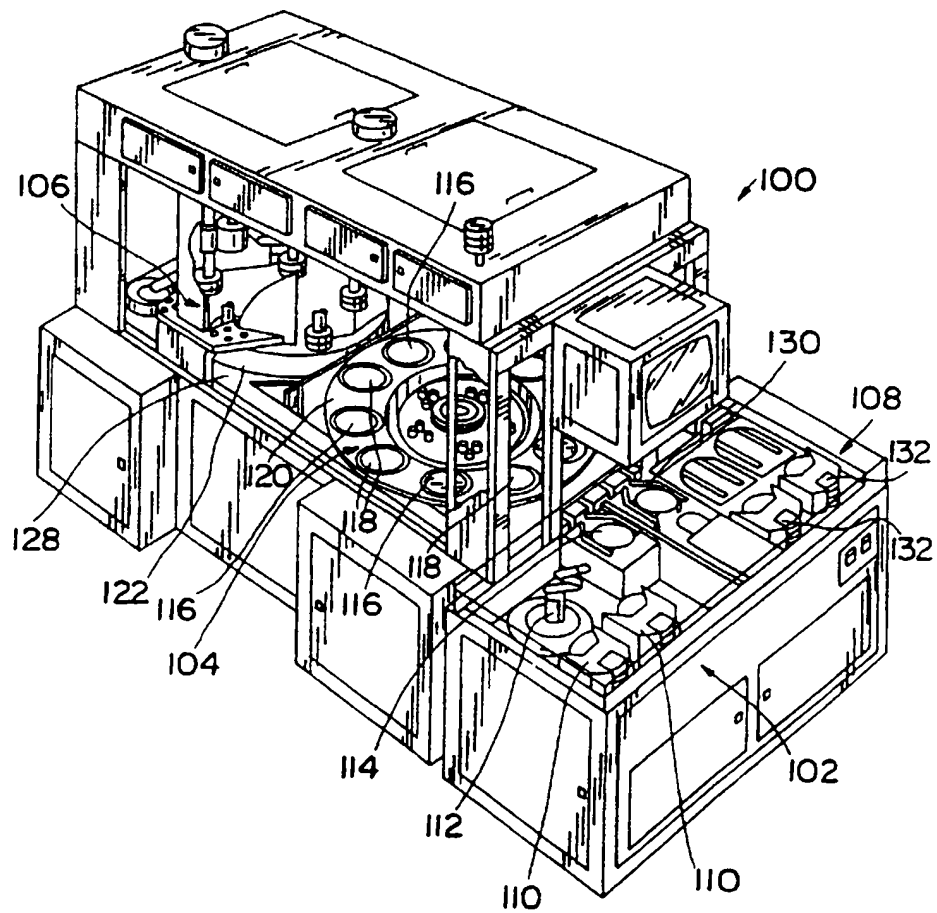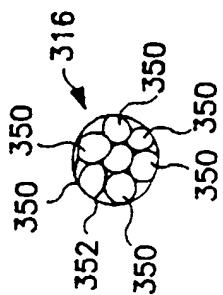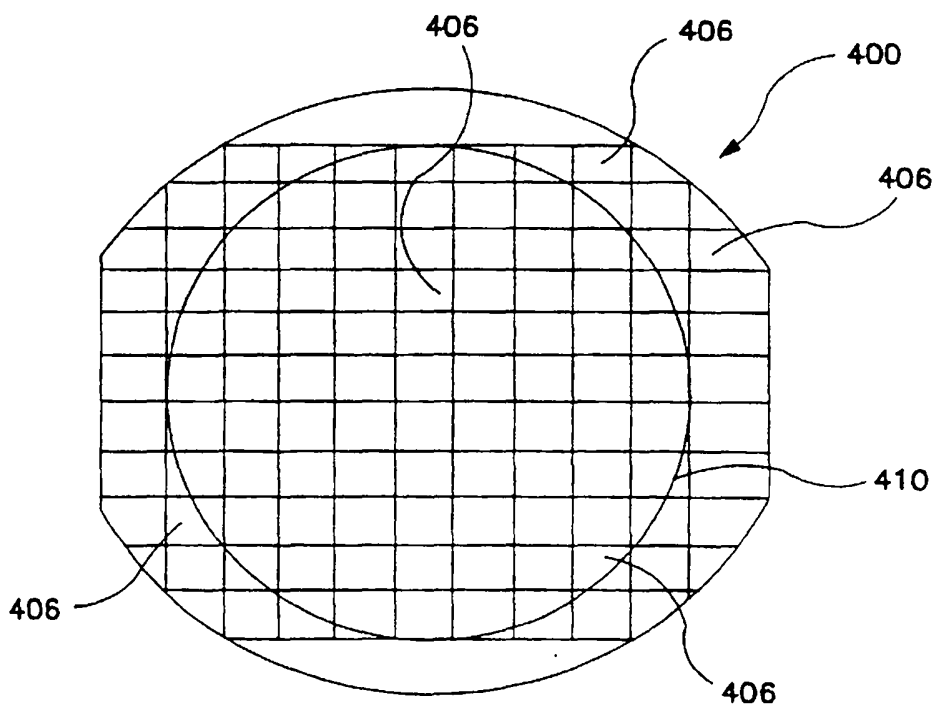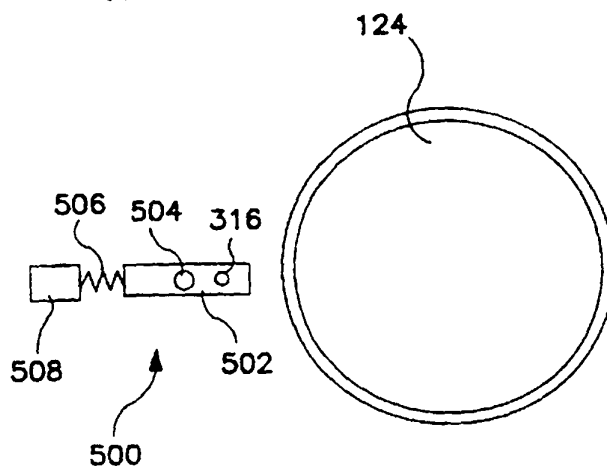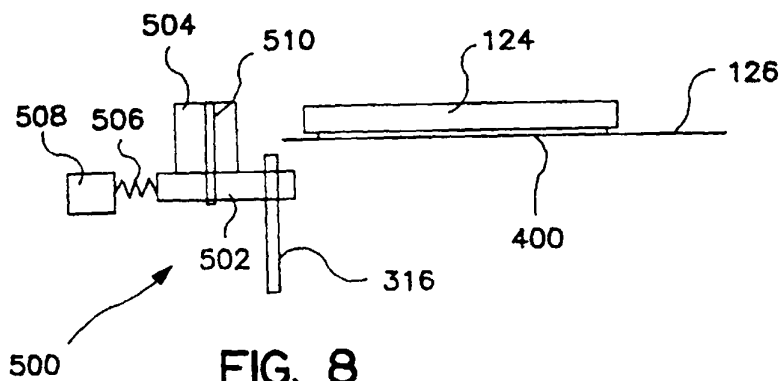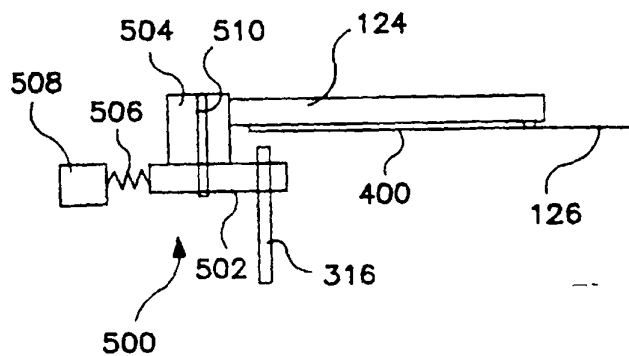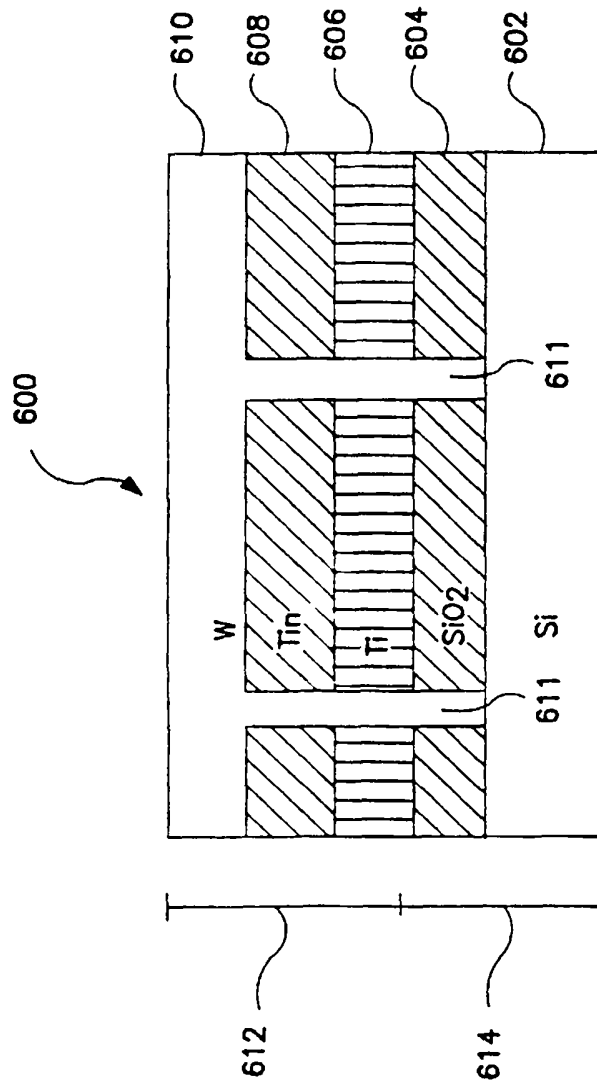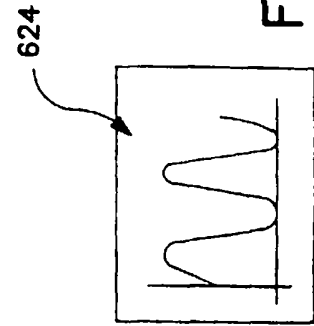
FIG. 1
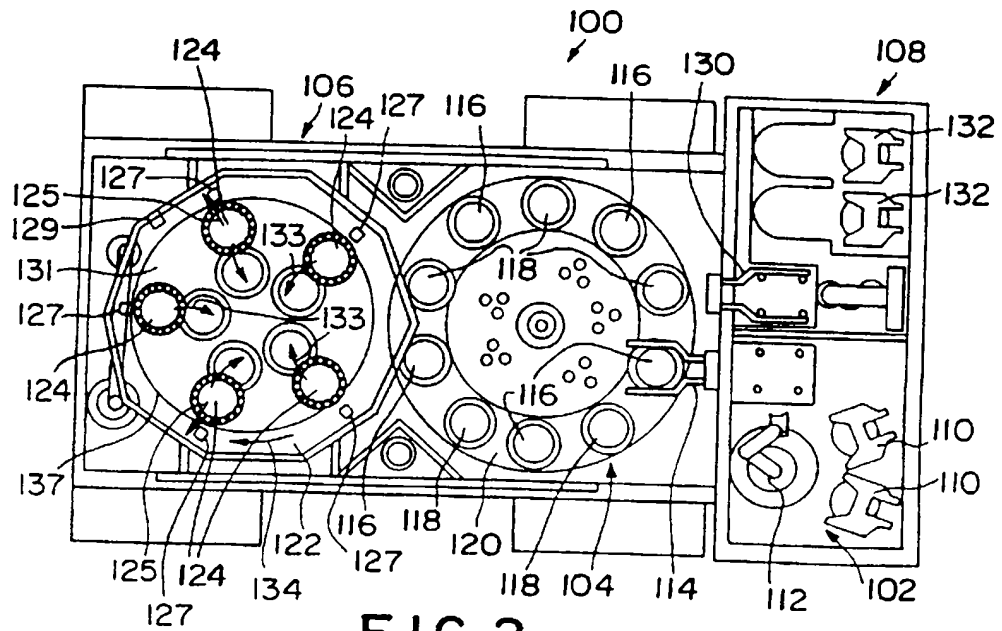
FIG. 2

3 / 8

FIG. 3

4 / 8
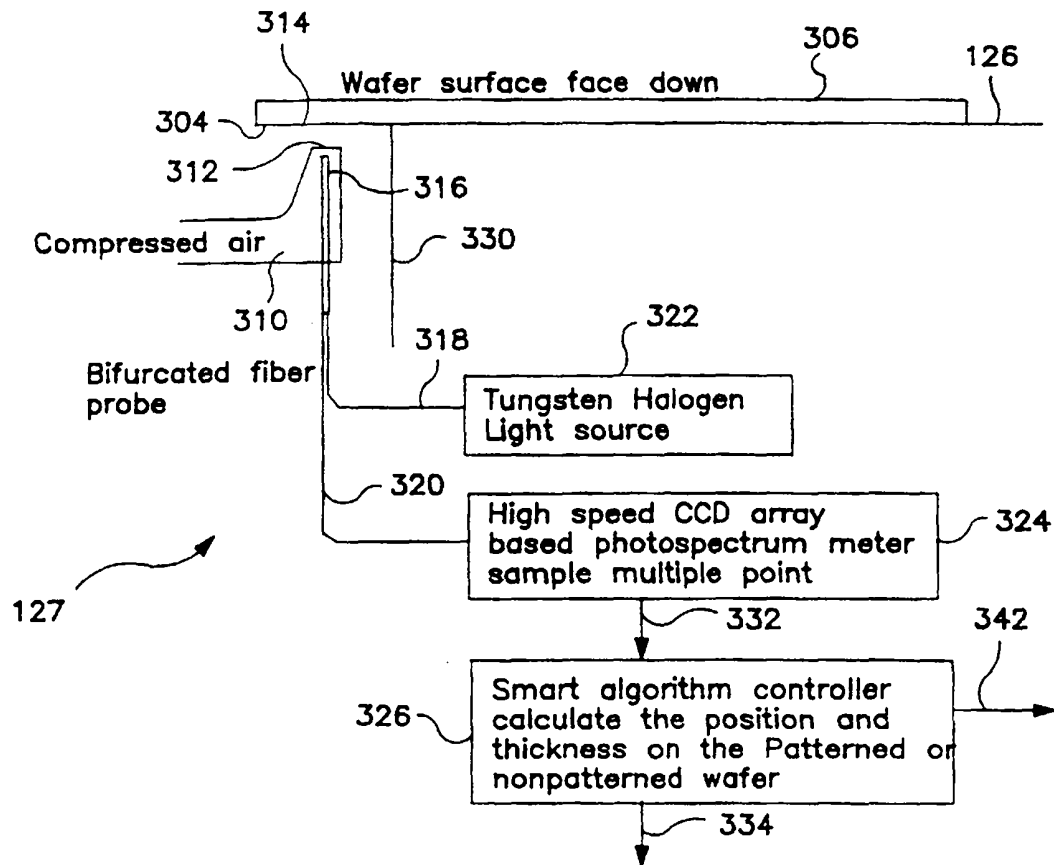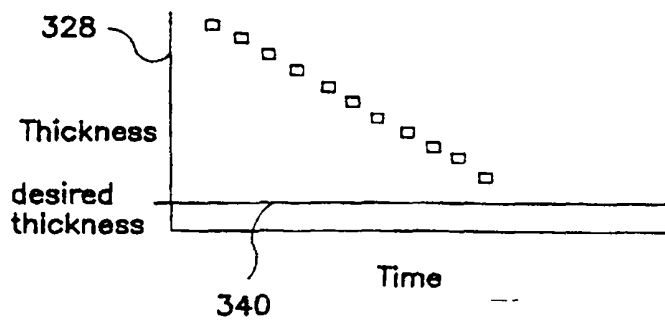
FIG. 5

FIG. 4
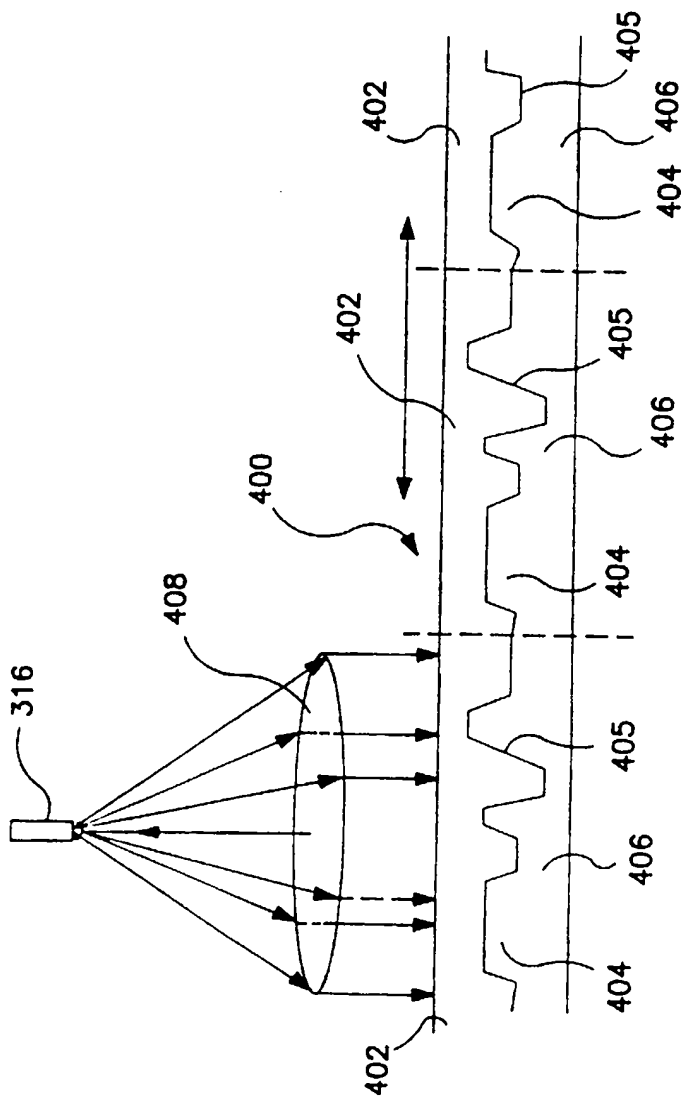
FIG. 6

FIG. 7

FIG. 8

FIG. 9

FIG. 10

FIG. 12

FIG. 14

FIG. 11

FIG. 13